



A Survey On Text Detection And Text Recognition Using Machine Learning Technique's

Vishnu Kant Soni¹, Dr. Vivek Shukla², Dr. S. R. Tandan³, Dr. Rohit Miri⁴

¹Research Scholar, Computer Science and Engineering, Dr. C. V. Raman University, Bilaspur C.G. India

²Assistant Professor, Computer Science and Engineering, Dr. C. V. Raman University, Bilaspur C.G., India

³Assistant Professor, Department of Computer Science, Government R.V.R.S. Kanya Mahavidyalaya, Kawardha, C.G, India.

⁴Associate Professor, Department of Computer Science and Engineering, UTD, CSVTU Bhilai, C.G. India

Abstract - Object detection methods based on machine learning are used in number of areas ITS (Intelligent Transportation System) and ADS (Autonomous Driving System). At the same time, the search for literature and popularity in special cases has also attracted a lot of attention and research. Content verification and authentication has become a major issue in the last few years. In computer vision and machine learning field, and the development of applications based on content detection and recognition, have also contributed to this trend. Searching for and familiarizing yourself with video subtitles and text on web pages is also helpful. Many studies have been done on text search and botanical image extraction. Various optical tools are also available. The problem of text finding and recognizing is still not fully solved. It is still very difficult to separate and extract texts from natural phenomena. Content analysis in natural pictures has attracted the devotion of the machine vision community due to its applications in data analysis, event data, robotic navigation, and image retrieval. Although previous studies have been quite successful in exploring the content and context of the text, it is still difficult due to the huge gap between the text and the date. The best method in text research is to divide the subject into content search and text recognition and treat these as separate tasks. Machine learning-based methods check for both.

Index Terms—Text detection, Text recognition, Segmentation, Text extraction, Natural language Processing, Machine Learning

I. INTRODUCTION

The operation of detection and recognition of text in images and videos are very difficult task, as evidenced by the appearance of the Robust Reading competitions [1][2][3] and the biennial ones in 2003, 2005, 2011, and 2013. International Workshop on Desk Analysis and Document Recognition (CBDAR), 2005-2013. The appearance of text translation applications on electronic devices, including iPhone and Android stages. Real-time translation into different languages has brought new attention to these issues. There are several key motives for this tendency, along with the need to increase the number of jobs. Text is the most communicative means of communication and may be inserted into data. This is possible in a way that presents it "observable" and/or understandable by others. The collection of huge amounts of "street view" data is like one application contributive to this [4][5]. This provides the ability to acquire and process images anytime, anywhere, making

text recognition convenient in a variety of environments. Third, advances in system vision and design recognition technologies have made solving complex problems more possible.

II. REVIEW OF LITERATURE

Many explorations have been expanded since so many last years. There were lots of measures were taken in account of learning methods in a couple of years, quite in fields like machine vision, text examination, object acknowledgment, and other data handling viewpoints.

The author describes the problem in segmentation of character and dependency in dictionary of text recognition in natural picture, and suggested a text recognition technique based on the attention method and CTC (Connection Time Classification) loss [6].

The researchers studied research papers on image text detection and recognition written by various authors from all over the world. All research work uses different technique and methods to identify text in images. Finally, we compared the accuracy, precision, and recall speed of different methods used in different research papers [7].

The author's worked in the rapidly growing field of computer, but the huge number of images and exponential growth of information made hard to filter and extract right information. This study focuses on whether research topics in English can be accurately defined, that is very valuable for improving the relevance of mining applications. In this study, examine how to use neural network technique to increases image text detection and recognition performance in backgrounds [8].

The author explains the moving tendencies of people using electronic media to consume data. Therefore, it's very easy method to consume data has created problematic called "fake news." This is a dangerous drawback of this disreputable term is that it affects persons views of the universe, current example of misrepresentation about the corona virus that has flooded the web. In today's world, it seems that 'information hygiene' is needed more than anything else. However, in practice, no practical solution can determine the

verification of the circulated contents. With an automated solution, our job provides a good solution for wrong news detection. The technique provides a set of selected features collective with term embeddings to expect fake or valid text [9].

The author explores a collection of images from smart devices in the outside world and then detects the edges of characters. It has been done after text detection and its analysis steps. The text has been recognized in two languages: Urdu and English. Text detection is a challenging works because text on signs will be blurry, will be broken ligatures, and variation in size etc. Text detection will be done by using the following equation: $Hr1 = \sum f_{i,j}$. MSER technique is used for text recognition of signs. It uses ANN, SVM classifiers, and HOG methods for classification and recognize text taken from natural landscapes or the outside world. Experiments done by this method on 500 images containing natural scenes, of which 70.2% of job were correctly recognized [10].

The author defines an Arabic character recognition system for Arabic text with neural network; a collection of instant-invariant descriptors and an artificial neural network (ANN) is used for classification of the text. Studies show that accuracy can reach 90% [11].

The authors define the data that has been classified into three sections. In 34,000 fonts, for training 70% are used, for testing 15% are used, and for the validation 15% are used. All feature selection methods are using for feature extraction, and forward and back propagation neural networks using for extraction. The method reached a most recognition rate 98.27% [12].

The researcher's explain oriented gradient using histogram (HOG) and artificial neural network (ANN) classifiers using feature extraction recognition. This method provides a high accuracy of 97.06% [13].

The researcher's worked on a novel method called Deep Text for scene area proposal text extraction in natural picture using completely convolutional neural networks (CNNs). First, suggest a new source area proposal network (Inception-RPN) that uses a multi-

scale window to move the source network over a convolutional feature map and concatenate the set of text features in front of the bounding box to produce high generatively. Second, proposed a robust detection system that integrates Ambiguous Text Category (ATC) contents and multilayer Area of Interest Fusion (MLRP) to increase text and non-text localization. This approach outperforms previous results, achieving F-scores of 0.83 and 0.85 on the 2011 and 2013 ICDAR strong text detection [14].

The author worked in the area of segmenting related characters in text CAPTCHA for intelligent character recognition. In this research, develop a machine recognition system for CAPTCHA. To achieve this goal, preprocessing is performed, followed by segmentation and recognition. The preprocessing step can be divided into four sub steps: grayscale conversion, binary conversion, cropping, and decimation. Experiments using were performed on CAPTCHA from various sources, including Taoba, MSN, and eBay, along with precisions of 51.3, 27.1, and 53.2, respectively [15].

The author detects and recognizes text from image of a natural scene is a difficult but very important work of finding information from the image. In present research, efficient method that detects extended maximum stable extreme areas (MSER) as important character candidates and filters these character applicants according to stroke width changes to remove areas were stroke width variations too much. To detect text areas, we first apply some preprocessing to the natural image, and then after MSER detection, we create the joining of sharp edges and MSER areas to detect areas that are more related belong to text. At last, the particular text zone is used as input too new OCR (Optical Character Recognition) method to make the text reusable. The experimentation results show that the f-measure of the ICDAR 2011 dataset resultant is 77.47% [16].

The researchers developed an Arabic text recognition technology by using Hare Cascade Classifier (HCC). The classification method trained and tested on with 2000 picture. Cascade boosting method is used for feature

extraction. This structure has been tested on actual images and gives 87% accuracy in recognizing Arabic characters [17].

The author developed SVM produces better results when we are using a linear kernel rather than Euclidean distances. Using a linear kernel gives a result of 95.86% [18].

The author developed the neural network system for Hindi optical character recognition and calculated using 1000 datasets. Video removal is based on a histogram measured for average distance, pixel value. A back propagation NN including two hidden layers was used for classify purpose. Test results showed that the recognition accuracy was 98.5% [19].

III. ANALYSIS

Analysis and comparison of various researches according to different strength and issues:

Table I. Different Methods Used for Text Detection with Parameters:

S.No	Methods	Strengths	Issues
1	Edges-based	Very Simple, fast, high recall	Conscious to complex background
2	Colour-based	Control many detection challenges e.g. Shadow effects	Not suited for low contrast font
3	Texture-based	Suited for poor resolution and low contrast text	Time taking, more complex in computation
4	CC-based	Simple, Suitable for text recognition	Issue in CC analysis
5	SWT	Efficient and stable	complex content increases complexity
6	MSER	Efficient, stable with good precision	Duplicate component generation, low recall
7	Geometric features	Simple, low cast and efficient	Maximum time use like supplementary features
8	HoG	Strongly describe text contour	Very sensitive to complex background

9	SVM	Best for classification like text / non-text	Computationally complex
10	Naïve Bayesian	Simple, fast	Highly Sensitive to noise
11	NN/Deep NN	Good results for text detection	Does not suitable for real-time applications

IV. CONCLUSION AND FUTURE WORKS

This article discussed various text detection and recognition algorithms. We compared various parameters such as intensity and challenge to get feedback on different types of text detection and text recognition algorithms. Researchers can use these facts to develop better algorithms. In the above study, there is no single algorithm that works in all noisy situations with multi-oriented text areas (blurry, dim, and shadowed). Our future work will build on the above results to develop more efficient algorithms that work in different types of environments with multi-oriented text areas.

REFERENCES

- [1] Lucas, S.M., Panaretos, A., Sosa, L., Tang, A., Wong, S., Young, R., Ashida, K., Nagai, H., Okamoto, M., Yamamoto, H. and Miyao, H., "ICDAR 2003 Robust Reading Competitions: Entries, Results, and Future Directions." International Journal of Document Analysis and Recognition (IJRAR), vol. 7, pp. 105-122, 2005.
- [2] Lucas, S.M, August. ICDAR "Text Locating Competition Results", In Eighth international conference on document analysis and recognition (ICDAR'05) pp. 80-84, 2005.
- [3] Shahab, A., Shafait, F. and Dengel, A, September. ICDAR "Robust Reading Competition Challenge 2: Reading Text in Scene Images", In 2011 international conference on document analysis and recognition pp. 1491-1496, 2011.
- [4] Chen, X., Yang, J., Zhang, J. and Waibel, A. "Automatic Detection and Recognition of Signs from Natural Scenes", IEEE Transactions on image processing, Vol. 13, no 1, pp. 87-99, 2004.
- [5] Chen, D., Odobez, J.M. and Boulard, H. "Text Detection and Recognition In Images And Video Frames", Pattern recognition, Vol. 37, no 3, pp. 595-608, 2004.
- [6] Jiang, Y., Jiang, Z., He, L. and Chen, S. "Text Recognition in Natural Scenes Based on Deep Learning", Multimedia Tools and Applications, 81(8), pp. 10545-10559, 2022.
- [7] Kulkarni, C.R. and Barbadekar, A.B. "Text Detection and Recognition: A Review", International Research Journal of Engineering and Technology (IRJET), Vol. 4, no. 6, pp. 179-185, 2017.
- [8] Li, C. "Research on Methods of English Text Detection and Recognition Based on Neural Network Detection Model", Scientific programming, 2021, pp.1-11, 2021.
- [9] Mertoğlu, U., Genç, B. and Sever, H. "Text-Based Fake News Detection via Machine Learning", In Trends in Data Engineering Methods for Intelligent Systems: Proceedings of the International Conference on Artificial Intelligence and Applied Mathematics in Engineering, pp. 113-124, 2020.
- [10] Panhwar, M.A., Memon, K.A., Abro, A., Zhongliang, D., Khuhro, S.A. and Memon, S. "Signboard Detection and Text Recognition Using Artificial Neural Networks", IEEE 9th international conference on electronics information and emergency communication (ICEIEC), pp. 16-19, July 2019.
- [11] Jebri, N.A., Al-Zoubi, H.R. and Abu Al-Haija, Q, "Recognition of Handwritten Arabic Characters Using Histograms of Oriented Gradient (HOG)", Pattern Recognition and Image Analysis, Vol. 28, pp. 321-345, 2018.
- [12] Lamghari, N., Charaf, M.E.H. and Raghay, S. "Hybrid Feature Vector for the Recognition of Arabic Handwritten Characters Using Feed-Forward Neural Network", Arabian Journal for Science and Engineering, Vol. 43, pp. 7031-7039, 2018.
- [13] Singh, N, February., "An Efficient Approach for Handwritten Devanagari Character Recognition Based on Artificial Neural Network", 5th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 894-897, 2018.

- [14] Zhong, Z., Jin, L. and Huang, S, March., “Deeptext: A New Approach for Text Proposal Generation and Text Detection in Natural Images”, IEEE international conference on acoustics, speech and signal processing (ICASSP), pp. 1208-1212, 2017.
- [15] Hussain, R., Gao, H. and Shaikh, R.A., “Segmentation of Connected Characters in Text-Based Captchas for Intelligent Character Recognition”, Multimedia Tools and Applications, Vol. 76, pp. 25547-25561, 2017.
- [16] Islam, M. R., Mondal, C., Azam, M. K. and Islam, A.S.M.J, “Text Detection and Recognition Using Enhanced MSER Detection”, 5th International Conference on Informatics, Electronics and Vision (ICIEV) pp. 15-20, 2016.
- [17] AbdelRaouf, A., Higgins, C. A., Pridmore, T. and Khalil, M. I., “Arabic Character Recognition Using a Haar Cascade Classifier Approach (HCC)”, Pattern analysis and applications, 19, pp. 411-426, 2016.
- [18] Gaur, A. and Yadav, S, January, “Handwritten Hindi Character Recognition Using K-Means Clustering and SVM”, 4th international symposium on emerging trends and technologies in libraries and information services. pp. 65-70, 2015.
- [19] Yadav, D., Sánchez-Cuadrado, S. and Morato, J., “Optical Character Recognition for Hindi Language Using a Neural-Network Approach”, Journal of Information Processing Systems, Vol. 9, no 1, pp. 117-140, 2013.

